A&A 616, A1 (2018)
https://doi.org/10.1051/0004-6361/201833051

Astronomy & Astrophysics

Special issue

Gaia Data Release 2

Gaia Data Release 2

Summary of the contents and survey properties

Gaia Collaboration, A. G. A. Brown[1,⋆], A. Vallenari[2], T. Prusti[3], J. H. J. de Bruijne[3], C. Babusiaux[4,5], C. A. L. Bailer-Jones[6], M. Biermann[7], D. W. Evans[8], L. Eyer[9], F. Jansen[10], C. Jordi[11], S. A. Klioner[12], U. Lammers[13], L. Lindegren[14], X. Luri[11], F. Mignard[15], C. Panem[16], D. Pourbaix[17,18], S. Randich[19], P. Sartoretti[4], H. I. Siddiqui[20], C. Soubiran[21], F. van Leeuwen[8], N. A. Walton[8], F. Arenou[4], U. Bastian[7], M. Cropper[22], R. Drimmel[23], D. Katz[4], M. G. Lattanzi[23], J. Bakker[13], C. Cacciari[24], J. Castañeda[11], L. Chaoul[16], N. Cheek[25], F. De Angeli[8], C. Fabricius[11], R. Guerra[13], B. Holl[9], E. Masana[11], R. Messineo[26], N. Mowlavi[9], K. Nienartowicz[27], P. Panuzzo[4], J. Portell[11], M. Riello[8], G. M. Seabroke[22], P. Tanga[15], F. Thévenin[15], G. Gracia-Abril[28,7], G. Comoretto[20], M. Garcia-Reinaldos[13], D. Teyssier[20], M. Altmann[7,29], R. Andrae[6], M. Audard[9], I. Bellas-Velidis[30], K. Benson[22], J. Berthier[31], R. Blomme[32], P. Burgess[8], G. Busso[8], B. Carry[15,31], A. Cellino[23], G. Clementini[24], M. Clotet[11], O. Creevey[15,33], M. Davidson[34], J. De Ridder[35], L. Delchambre[36], A. Dell'Oro[19], C. Ducourant[21], J. Fernández-Hernández[37], M. Fouesneau[6], Y. Frémat[32], L. Galluccio[15], M. García-Torres[38], J. González-Núñez[25,39], J. J. González-Vidal[11], E. Gosset[36,18], L. P. Guy[27,40], J.-L. Halbwachs[41], N. C. Hambly[34], D. L. Harrison[8,42], J. Hernández[13], D. Hestroffer[31], S. T. Hodgkin[8], A. Hutton[43], G. Jasniewicz[44], A. Jean-Antoine-Piccolo[16], S. Jordan[7], A. J. Korn[45], A. Krone-Martins[46], A. C. Lanzafame[47,48], T. Lebzelter[49], W. Löffler[7], M. Manteiga[50,51], P. M. Marrese[52,53], J. M. Martín-Fleitas[43], A. Moitinho[46], A. Mora[43], K. Muinonen[54,55], J. Osinde[56], E. Pancino[19,53], T. Pauwels[32], J.-M. Petit[57], A. Recio-Blanco[15], P. J. Richards[58], L. Rimoldini[27], A. C. Robin[57], L. M. Sarro[59], C. Siopis[17], M. Smith[22], A. Sozzetti[23], M. Süveges[6], J. Torra[11], W. van Reeven[43], U. Abbas[23], A. Abreu Aramburu[60], S. Accart[61], C. Aerts[35,62], G. Altavilla[52,53,24], M. A. Álvarez[50], R. Alvarez[13], J. Alves[49], R. I. Anderson[63,9], A. H. Andrei[64,65,29], E. Anglada Varela[37], E. Antiche[11], T. Antoja[3,11], B. Arcay[50], T. L. Astraatmadja[6,66], N. Bach[43], S. G. Baker[22], L. Balaguer-Núñez[11], P. Balm[20], C. Barache[29], C. Barata[46], D. Barbato[67,23], F. Barblan[9], P. S. Barklem[45], D. Barrado[68], M. Barros[46], M. A. Barstow[69], S. Bartholomé Muñoz[11], J.-L. Bassilana[61], U. Becciani[48], M. Bellazzini[24], A. Berihuete[70], S. Bertone[23,29,71], L. Bianchi[72], O. Bienaymé[41], S. Blanco-Cuaresma[9,21,73], T. Boch[41], C. Boeche[2], A. Bombrun[74], R. Borrachero[11], D. Bossini[2], S. Bouquillon[29], G. Bourda[21], A. Bragaglia[24], L. Bramante[26], M. A. Breddels[75], A. Bressan[76], N. Brouillet[21], T. Brüsemeister[7], E. Brugaletta[48], B. Bucciarelli[23], A. Burlacu[16], D. Busonero[23], A. G. Butkevich[12], R. Buzzi[23], E. Caffau[4], R. Cancelliere[77], G. Cannizzaro[78,62], T. Cantat-Gaudin[2,11], R. Carballo[79], T. Carlucci[29], J. M. Carrasco[11], L. Casamiquela[11], M. Castellani[52], A. Castro-Ginard[11], P. Charlot[21], L. Chemin[80], A. Chiavassa[15], G. Cocozza[24], G. Costigan[1], S. Cowell[8], F. Crifo[4], M. Crosta[23], C. Crowley[74], J. Cuypers[†32], C. Dafonte[50], Y. Damerdji[36,81], A. Dapergolas[30], P. David[31], M. David[82], P. de Laverny[15], F. De Luise[83], R. De March[26], D. de Martino[84], R. de Souza[85], A. de Torres[74], J. Debosscher[35], E. del Pozo[43], M. Delbo[15], A. Delgado[8], H. E. Delgado[59], P. Di Matteo[4], S. Diakite[57], C. Diener[8], E. Distefano[48], C. Dolding[22], P. Drazinos[86], J. Durán[56], B. Edvardsson[45], H. Enke[87], K. Eriksson[45], P. Esquej[88], G. Eynard Bontemps[16], C. Fabre[89], M. Fabrizio[52,53], S. Faigler[90], A. J. Falcão[91], M. Farràs Casas[11], L. Federici[24], G. Fedorets[54], P. Fernique[41], F. Figueras[11], F. Filippi[26], K. Findeisen[4], A. Fonti[26], E. Fraile[88], M. Fraser[8,92], B. Frézouls[16], M. Gai[23], S. Galleti[24], D. Garabato[50], F. García-Sedano[59], A. Garofalo[93,24], N. Garralda[11], A. Gavel[45], P. Gavras[4,30,86], J. Gerssen[87], R. Geyer[12], P. Giacobbe[23], G. Gilmore[8], S. Girona[94], G. Giuffrida[53,52], F. Glass[9], M. Gomes[46], M. Granvik[54,95], A. Gueguen[4,96], A. Guerrier[61], J. Guiraud[16], R. Gutiérrez-Sánchez[20], R. Haigron[4], D. Hatzidimitriou[86,30], M. Hauser[7,6], M. Haywood[4], U. Heiter[45], A. Helmi[75], J. Heu[4], T. Hilger[12], D. Hobbs[14], W. Hofmann[7], G. Holland[8], H. E. Huckle[22], A. Hypki[1,97], V. Icardi[26], K. Janßen[87], G. Jevardat de Fombelle[27], P. G. Jonker[78,62], Á. L. Juhász[98,99], F. Julbe[11], A. Karampelas[86,100], A. Kewley[8], J. Klar[87], A. Kochoska[101,102]

⋆ Corresponding author: A. G. A. Brown, e-mail: brown@strw.leidenuniv.nl

R. Kohley[13], K. Kolenberg[103,35,73], M. Kontizas[86], E. Kontizas[30], S. E. Koposov[8,104], G. Kordopatis[15], Z. Kostrzewa-Rutkowska[78,62], P. Koubsky[105], S. Lambert[29], A. F. Lanza[48], Y. Lasne[61], J.-B. Lavigne[61], Y. Le Fustec[106], C. Le Poncin-Lafitte[29], Y. Lebreton[4,107], S. Leccia[84], N. Leclerc[4], I. Lecoeur-Taibi[27], H. Lenhardt[7], F. Leroux[61], S. Liao[23,108,109], E. Licata[72], H. E. P. Lindstrøm[110,111], T. A. Lister[112], E. Livanou[86], A. Lobel[32], M. López[68], S. Managau[61], R. G. Mann[34], G. Mantelet[7], O. Marchal[4], J. M. Marchant[113], M. Marconi[84], S. Marinoni[52,53], G. Marschalkó[98,114], D. J. Marshall[115], M. Martino[26], G. Marton[98], N. Mary[61], D. Massari[75], G. Matijevič[87], T. Mazeh[90], P. J. McMillan[14], S. Messina[48], D. Michalik[14], N. R. Millar[8], D. Molina[11], R. Molinaro[84], L. Molnár[98], P. Montegriffo[24], R. Mor[11], R. Morbidelli[23], T. Morel[36], D. Morris[34], A. F. Mulone[26], T. Muraveva[24], I. Musella[84], G. Nelemans[62,35], L. Nicastro[24], L. Noval[6], W. O'Mullane[13,40], C. Ordénovic[15], D. Ordóñez-Blanco[27], P. Osborne[8], C. Pagani[69], I. Pagano[48], F. Pailler[16], H. Palacin[61], L. Palaversa[8,9], A. Panahi[90], M. Pawlak[116,117], A. M. Piersimoni[83], F.-X. Pineau[41], E. Plachy[98], G. Plum[4], E. Poggio[67,23], E. Poujoulet[118], A. Prša[102], L. Pulone[52], E. Racero[25], S. Ragaini[24], N. Rambaux[31], M. Ramos-Lerate[119], S. Regibo[35], C. Reylé[57], F. Riclet[16], V. Ripepi[84], A. Riva[23], A. Rivard[61], G. Rixon[8], T. Roegiers[120], M. Roelens[9], M. Romero-Gómez[11], N. Rowell[34], F. Royer[4], L. Ruiz-Dern[4], G. Sadowski[17], T. Sagristà Sellés[7], J. Sahlmann[13,121], J. Salgado[122], E. Salguero[37], N. Sanna[19], T. Santana-Ros[97], M. Sarasso[23], H. Savietto[123], M. Schultheis[15], E. Sciacca[48], M. Segol[124], J. C. Segovia[25], D. Ségransan[9], I-C. Shih[4], L. Siltala[54,125], A. F. Silva[46], R. L. Smart[23], K. W. Smith[6], E. Solano[68,126], F. Solitro[26], R. Sordo[2], S. Soria Nieto[11], J. Souchay[29], A. Spagna[23], F. Spoto[15,31], U. Stampa[7], I. A. Steele[113], H. Steidelmüller[12], C. A. Stephenson[20], H. Stoev[127], F. F. Suess[8], J. Surdej[36], L. Szabados[98], E. Szegedi-Elek[98], D. Tapiador[128,129], F. Taris[29], G. Tauran[61], M. B. Taylor[130], R. Teixeira[85], D. Terrett[58], P. Teyssandier[29], W. Thuillot[31], A. Titarenko[15], F. Torra Clotet[131], C. Turon[4], A. Ulla[132], E. Utrilla[43], S. Uzzi[26], M. Vaillant[61], G. Valentini[83], V. Valette[16], A. van Elteren[1], E. Van Hemelryck[32], M. van Leeuwen[8], M. Vaschetto[26], A. Vecchiato[23], J. Veljanoski[75], Y. Viala[4], D. Vicente[94], S. Vogt[120], C. von Essen[133], H. Voss[11], V. Votruba[105], S. Voutsinas[34], G. Walmsley[16], M. Weiler[11], O. Wertz[134], T. Wevers[8,62], Ł. Wyrzykowski[8,116], A. Yoldas[8], M. Žerjal[101,135], H. Ziaeepour[57], J. Zorec[136], S. Zschocke[12], S. Zucker[137], C. Zurbach[44], T. Zwitter[101]

(Affiliations can be found after the references)

Received 19 March 2018 / Accepted 14 April 2018

ABSTRACT

Context. We present the second *Gaia* data release, *Gaia* DR2, consisting of astrometry, photometry, radial velocities, and information on astrophysical parameters and variability, for sources brighter than magnitude 21. In addition epoch astrometry and photometry are provided for a modest sample of minor planets in the solar system.

Aims. A summary of the contents of *Gaia* DR2 is presented, accompanied by a discussion on the differences with respect to *Gaia* DR1 and an overview of the main limitations which are still present in the survey. Recommendations are made on the responsible use of *Gaia* DR2 results.

Methods. The raw data collected with the *Gaia* instruments during the first 22 months of the mission have been processed by the *Gaia* Data Processing and Analysis Consortium (DPAC) and turned into this second data release, which represents a major advance with respect to *Gaia* DR1 in terms of completeness, performance, and richness of the data products.

Results. *Gaia* DR2 contains celestial positions and the apparent brightness in G for approximately 1.7 billion sources. For 1.3 billion of those sources, parallaxes and proper motions are in addition available. The sample of sources for which variability information is provided is expanded to 0.5 million stars. This data release contains four new elements: broad-band colour information in the form of the apparent brightness in the $G_{\rm BP}$ (330–680 nm) and $G_{\rm RP}$ (630–1050 nm) bands is available for 1.4 billion sources; median radial velocities for some 7 million sources are presented; for between 77 and 161 million sources estimates are provided of the stellar effective temperature, extinction, reddening, and radius and luminosity; and for a pre-selected list of 14 000 minor planets in the solar system epoch astrometry and photometry are presented. Finally, *Gaia* DR2 also represents a new materialisation of the celestial reference frame in the optical, the *Gaia*-CRF2, which is the first optical reference frame based solely on extragalactic sources. There are notable changes in the photometric system and the catalogue source list with respect to *Gaia* DR1, and we stress the need to consider the two data releases as independent.

Conclusions. *Gaia* DR2 represents a major achievement for the *Gaia* mission, delivering on the long standing promise to provide parallaxes and proper motions for over 1 billion stars, and representing a first step in the availability of complementary radial velocity and source astrophysical information for a sample of stars in the *Gaia* survey which covers a very substantial fraction of the volume of our galaxy.

Key words. catalogs – astrometry – techniques: radial velocities – stars: fundamental parameters – stars: variables: general – minor planets, asteroids: general

1. Introduction

We present the second intermediate *Gaia* data release (*Gaia* Data Release 2, *Gaia* DR2), which is based on the data collected during the first 22 months of the nominal mission lifetime (scientific data collection started in July 2014 and nominally lasts 60 months, see Gaia Collaboration 2016b). *Gaia* DR2 represents the planned major advance with respect to the first intermediate *Gaia* data release (*Gaia* DR1, Gaia Collaboration 2016a), making the leap to a high-precision parallax and proper motion catalogue for over 1 billion sources, supplemented by precise and homogeneous multi-band all-sky photometry and a large radial velocity survey at the bright ($G \lesssim 13$) end. The availability of precise fundamental astrophysical information required to map and understand the Milky Way is thus expanded to a very substantial fraction of the volume of our galaxy, well beyond the immediate solar neighbourhood. The data diversity of *Gaia* DR2 is also significantly enhanced with respect to *Gaia* DR1 through the availability of astrophysical parameters for a large sample of stars, the significant increase in the number and types of variable stars and their light curves, and the addition for the first time of solar system astrometry and photometry. This paper is structured as follows. In Sect. 2 we provide a short overview of the improvements and additions to the data processing that led to the production of *Gaia* DR2. We summarise the contents of the second data release in Sect. 3 and illustrate the quality of this release through all-sky maps of source counts and colours in Sect. 4. In Sect. 5 we discuss the major differences between *Gaia* DR2 and *Gaia* DR1, in particular pointing out the evolution of the source list and the need to always qualify *Gaia* source identifiers with the data release they refer to. The two releases should be treated as entirely independent catalogues. The known limitations of the second *Gaia* data release are presented in Sect. 6 and additional guidance on the use of the data is provided in Sect. 7. In Sect. 8 we provide updates to the *Gaia* data access facilities and documentation available to the astronomical community. We conclude with a look ahead at the next release in Sect. 9. Throughout the paper we make reference to other DPAC papers that provide more details on the data processing and validation for *Gaia* DR2. All these papers (together with the present article) can be found in the Astronomy & Astrophysics Special edition on *Gaia* DR2.

2. Data processing for *Gaia* DR2

To provide the context for the description of the data release contents in the next section, we provide here a summary of the input measurements used and the main additions and improvements implemented in the data processing for *Gaia* DR2. We recall that *Gaia* measurements are collected with three instruments. The astrometric instrument collects images in *Gaia*'s white-light *G*-band (330–1050 nm); the Blue (BP) and Red (RP) prism photometers collect low resolution spectrophotometric measurements of source spectral energy distributions over the wavelength ranges 330–680 nm and 630–1050 nm, respectively; and the radial velocity spectrometer (RVS) collects medium resolution ($R \sim 11\,700$) spectra over the wavelength range 845–872 nm centred on the Calcium triplet region. For more details on the *Gaia* instruments and measurements we refer to Gaia Collaboration (2016b). The RVS, from which results are presented in *Gaia* DR2 for the first time, is described in detail in Cropper et al. (2018). An important part of the pre-processing for all *Gaia* instruments is to remove the effect of non-uniformity of the CCD bias levels, which is essential for achieving the ultimate image location and radial velocity determination performance. The details of this process are described in Hambly et al. (2018).

The timing of events on board *Gaia*, including the data collection, is given in terms of the on board mission timeline (OBMT) which is generated by the *Gaia* on board clock. By convention OBMT is expressed in units of 6 h (21 600 s) spacecraft revolutions (Gaia Collaboration 2016b). The approximate relation between OBMT (in revolutions) and the barycentric coordinate time (TCB, in Julian years) at *Gaia* is

$$\mathrm{TCB} \simeq \mathrm{J2015.0} + (\mathrm{OBMT} - 1717.6256\ \mathrm{rev})/(1461\ \mathrm{rev\ yr^{-1}})\,. \quad (1)$$

The 22 month time interval covered by the observations used for *Gaia* DR2 starts at OBMT 1078.3795 rev = J2014.5624599 TCB (approximately 2014 July 25, 10:30:00 UTC), and ends at OBMT 3750.5602 rev = J2016.3914678 TCB (approximately 2016 May 23, 11:35:00 UTC). As discussed in Gaia Collaboration (2016a) this time interval contains gaps caused by both spacecraft events and by on-ground data processing problems. This leads to gaps in the data collection or stretches of time over which the input data cannot be used. Which data are considered unusable varies across the *Gaia* data processing systems (astrometry, photometry, etc) and as a consequence the effective amount of input data used differs from one system to the other. We refer to the specific data processing papers (listed below) for the details.

A broad overview of the data processing for *Gaia* is given in Gaia Collaboration (2016b) while the simplified processing for *Gaia* DR1 is summarised in Gaia Collaboration (2016a), in particular in their Fig. 10. With respect to *Gaia* DR1 the following major improvements were implemented in the astrometric processing (for details, see Lindegren et al. 2018):

- Creation of the source list: this process (also known as cross-matching; Fabricius et al. 2016) provides the link between the individual *Gaia* detections and the entries ("sources") in the *Gaia* working catalogue. For *Gaia* DR1 the detections were matched to the nearest source, using a match radius of 1.5 arcsec, and new sources were created when no match was found. Spurious detections and limitations of the initial source list resulted in many spurious sources but also the loss in *Gaia* DR1 of many real sources, including high proper motion stars. For *Gaia* DR2 the source list was created essentially from scratch, based directly on the detections and using a cluster analysis algorithm that takes into account a possible linear motion of the source. The source list for *Gaia* DR2 is therefore much cleaner and of higher angular resolution (Sect. 5.3), resulting in improved astrometry.
- Attitude modelling: in the astrometric solution, the pointing of the instrument is modelled as a function of time using splines. However, these cannot represent rapid variations caused by the active attitude control, micro-clanks (microscopic structural changes in the spacecraft), and micrometeoroid hits. In *Gaia* DR1 the accuracy of the attitude determination was limited by such effects. For *Gaia* DR2 the rapid variations are determined and subtracted by a dedicated process, using rate measurements from successive CCD observations of bright sources.
- Calibration modelling: optical aberrations in the telescopes and the wavelength-dependent diffraction create colour-dependent shifts of the stellar images (chromaticity). This will eventually be handled in the pre-processing of the raw data, by fitting colour-dependent PSFs or LSFs to the CCD samples. This procedure will only be in place for the next

release, and the effect was completely ignored for *Gaia* DR1. In the current astrometric solution chromaticity is handled by the introduction of colour-dependent terms in the geometric calibration model.
– Global modelling: the basic-angle variations are more accurately modelled thanks to an improved processing of the on-board measurements (using the Basic Angle Monitor) and the introduction of global corrections to these measurements as additional unknowns in the astrometric solution. This has been especially important for reducing large-scale systematics in the parallaxes.
– Celestial reference frame: establishing a link to the extragalactic reference frame was complicated and indirect in *Gaia* DR1, which relied on the HIPPARCOS and *Tycho*-2 catalogues for the determination of proper motions. By contrast, *Gaia* DR2 contains the positions and proper motions for about half a million identified quasars, which directly define a very accurate celestial reference frame (*Gaia*-CRF2), as described in Gaia Collaboration (2018e).

The various improvements in the astrometric models have reduced the RMS residual of typical observations of bright stars ($G \lesssim 13$) from about 0.67 mas in *Gaia* DR1 to 0.2–0.3 mas in *Gaia* DR2.

Additional improvements in the data processing for *Gaia* DR2 as well as the introduction of new elements facilitated the much expanded variety of data published in this second release. Although the photometric processing pipeline did treat the data from *Gaia*'s BP and RP photometers from the start of the mission operations, it was decided not to publish the results in *Gaia* DR1 (Evans et al. 2017) because of the still preliminary nature of the calibrations of these instruments. The processing for *Gaia* DR2 features enhancements in the photometric calibrations, including of the BP and RP prism spectra. The integrated light from these spectra is published in this release as the fluxes in the $G_{\rm BP}$ and $G_{\rm RP}$ passbands. In addition the photometric passbands for G, $G_{\rm BP}$, and $G_{\rm RP}$ are published, both the versions used in the data processing and the revised versions (based on a deeper analysis involving the BP/RP spectra of standard stars). The photometric data processing and results validation for *Gaia* DR2 are described in Evans et al. (2018) and Riello et al. (2018).

The processing of RVS data was also in place from the start of mission operations but during the operations up to *Gaia* DR1 the adaptations necessary to the RVS pipeline to deal with the effects of the excess stray light on board *Gaia* prevented the publication of results. Hence *Gaia* DR2 features the first RVS results in the form of median radial velocities. The details of the RVS data processing and results validation are provided in Sartoretti et al. (2018), Katz et al. (2018), and Soubiran et al. (2018).

Epoch astrometry was determined for a list of 14 000 pre-selected small solar system bodies (henceforth referred to as Solar System Objects or SSOs). The data processing and validation for the *Gaia* DR2 SSO data are described in Gaia Collaboration (2018f).

Astrophysical parameters ($T_{\rm eff}$, A_G, $E(G_{\rm BP} - G_{\rm RP})$, radius and luminosity) were determined for between 77 and 161 million stars from the *Gaia* broad-band photometry and parallaxes alone (no non-*Gaia* data was used). The details of the astrophysical parameter estimation and the validation of the results are described in Andrae et al. (2018).

Practically all sources present in *Gaia* DR2 were analysed for apparent brightness variations, resulting in a catalogue of about 0.5 million stars securely identified as variables and for which light curves and statistical information on the photometric time series are provided. The variability processing is described in Holl et al. (2018).

Finally, an overall validation of the *Gaia* DR2 catalogue is described in Arenou et al. (2018), which, as outlined in Gaia Collaboration (2016b), involves an extensive scientific validation of the combined data presented in this data release.

A number of important shortcomings remain in the data processing, leading to limitations in *Gaia* DR2 which require taking some care when using the data. In Sect. 6 we summarise the known limitations of the present *Gaia* data release and point out, where relevant, the causes. Section 7 provides additional guidance on the use of *Gaia* DR2 results. The reader is strongly encouraged to read the papers listed above and the online documentation[1] to understand the limitations in detail.

Table 1. Number of sources of a given type or the number for which a given data product is available in *Gaia* DR2.

Data product or source type	Number of sources
Total	1 692 919 135
5-parameter astrometry	1 331 909 727
2-parameter astrometry	361 009 408
ICRF3 prototype sources	2820
Gaia-CRF2 sources	556 869
G-band	1 692 919 135
$G_{\rm BP}$-band	1 381 964 755
$G_{\rm RP}$-band	1 383 551 713
Radial velocity	7 224 631
Classified as variable	550 737
Variable type estimated	363 969
Detailed characterisation of light curve	390 529
Effective temperature $T_{\rm eff}$	161 497 595
Extinction A_G	87 733 672
Colour excess $E(G_{\rm BP} - G_{\rm RP})$	87 733 672
Radius	76 956 778
Luminosity	76 956 778
SSO epoch astrometry and photometry	14 099

3. Overview of the contents of *Gaia* DR2

Gaia DR2 contains astrometry, broad-band photometry, radial velocities, variable star classifications as well as the characterisation of the corresponding light curves, and astrophysical parameter estimates for a total of 1 692 919 135 sources. In addition the epoch astrometry and photometry for 14 099 solar system objects are listed. Basic statistics on the source numbers and the overall distribution in G can be found in Table 1 and Table 2, where it should noted that 4 per cent of the sources are fainter than $G = 21$. The overall quality of *Gaia* DR2 results in terms of the typically achieved uncertainties is summarised in Table 3. The contents of the main components of the release, of which the magnitude distributions are shown in Figs. 1 and 2, are summarised in the following paragraphs. We defer the discussion on the known limitations of *Gaia* DR2 to Sect. 6.

[1] http://gea.esac.esa.int/archive/documentation/GDR2/index.html

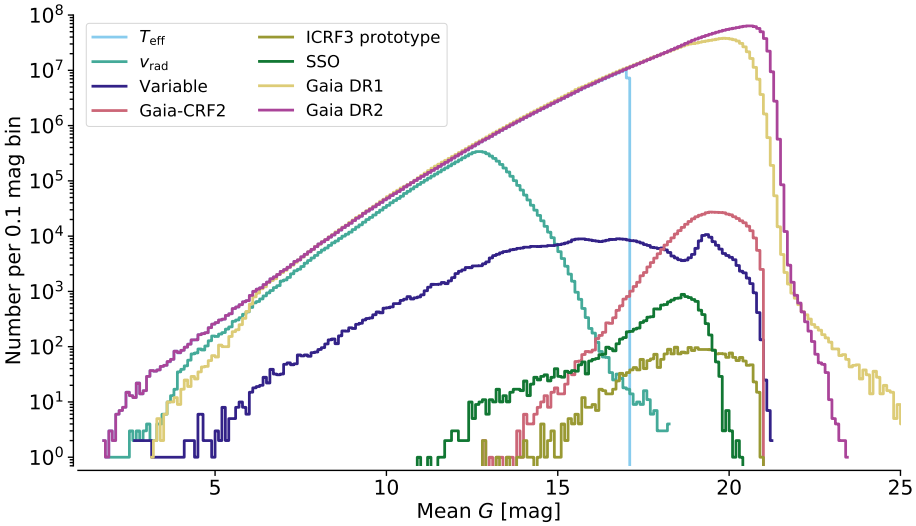

Fig. 1. Distribution of the mean values of G for all *Gaia* DR2 sources shown as histograms with 0.1 mag wide bins. The distribution of the *Gaia* DR1 sources is included for comparison and illustrates the improved photometry at the faint end and the improved completeness at the bright end. The other histograms are for the main *Gaia* DR2 components as indicated in the legend. See text for further explanations on the characteristics of the histograms.

Table 2. Distribution of the *Gaia* DR2 sources in G-band magnitude.

	Magnitude distribution percentiles (G)		
Percentile	All	5-parameter	2-parameter
0.135%	11.6	11.4	15.3
2.275%	15.0	14.7	18.5
15.866%	17.8	17.4	19.8
50%	19.6	19.3	20.6
84.134%	20.6	20.3	21.0
97.725%	21.1	20.8	21.2
99.865%	21.3	20.9	21.4

Notes. The distribution percentiles are shown for all sources and for those with a 5-parameter and 2-parameter astrometric solution, respectively.

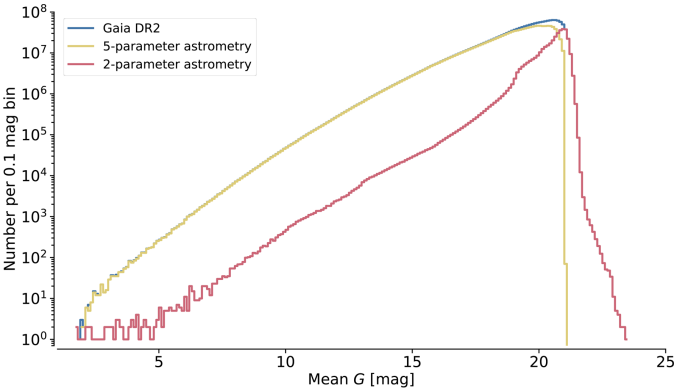

Fig. 2. Distribution of the mean values of G for the sources with a full astrometric solution in *Gaia* DR2 ("5-parameter") and for the sources for which only the celestial position is listed ("2-parameter") compared to the overall magnitude distribution for *Gaia* DR2.

3.1. Astrometric data set

The astrometric data set consists of two subsets: for 1 331 909 727 sources the full five-parameter astrometric solution is provided ("5-parameter" in Table 1), hence including celestial position, parallax, and proper motion. For the remaining 361 009 408 sources ("2-parameter" in Table 1) only the celestial positions (α, δ) are reported. Figure 2 shows the distribution in G for the 5-parameter and 2-parameter sources compared to the overall magnitude distribution. The 2-parameter sources are typically faint (with about half those sources at $G > 20.6$, see Table 2), have very few observations, or very poorly fit the five-parameter astrometric model. All sources fainter than $G = 21$ have only positions in *Gaia* DR2. We refer to Lindegren et al. (2018) for the detailed criteria used during the data processing to decide which type of solution should be adopted.

For a 2-parameter source the position was computed using a special fall-back solution. Rather than ignoring the parallax and proper motion of the source (i.e. assuming that they are strictly zero), the fall-back solution estimates all five parameters but applies a prior that effectively constrains the parallax and proper motion to realistically small values, depending on the magnitude and Galactic coordinates of the source (Michalik et al. 2015b). The resulting position is usually more precise, and its uncertainty more realistic (larger), than if only the position had been solved for. The parallax and proper motion of the fall-back solution may however be strongly biased, which is why they are not published.

The reference epoch for all (5- and 2-parameter) sources is J2015.5 (TCB). This epoch, close to the mid-time of the observations included in *Gaia* DR2, was chosen to minimise correlations between the position and proper motion parameters. This epoch is 0.5 yr later than the reference epoch for *Gaia* DR1, which must be taken into account when comparing the positions between the two releases.

As for *Gaia* DR1 all sources were treated as single stars when solving for the astrometric parameters. For a binary the parameters may thus refer to either component, or to the photocentre of the system, and the proper motion represents the mean motion of the component, or photocentre, over the 1.75 yr of data included in the solution. Depending on the orbital motion, this could be significantly different from the proper motion of the same object in *Gaia* DR1 (see Sect. 5).

The positions and proper motions are given in the second realisation of the *Gaia* celestial reference frame (*Gaia*-CRF2) which at the faint end ($G \sim 19$) is aligned with the International Celestial Reference Frame (ICRF) to about 0.02 mas RMS at epoch J2015.5 (TCB), and non-rotating with respect to the ICRF to within 0.02 mas yr^{-1}RMS. At the bright end ($G < 12$) the alignment can only be confirmed to be better than 0.3 mas while the bright reference frame is non-rotating to within 0.15 mas yr^{-1}. For details we refer to Lindegren et al. (2018). The *Gaia*-CRF2 is materialised by 556 869 QSOs and aligned

to the forthcoming version 3 of the ICRF through a subset of 2820 QSOs. It represents the first ever optical reference frame constructed on the basis of extragalactic sources only. The construction and properties of the *Gaia*-CRF2 as well as the comparison to the ICRF3 prototype are described in Gaia Collaboration (2018e).

3.2. Photometric data set

The photometric data set contains the broad band photometry in the G, $G_{\rm BP}$, and $G_{\rm RP}$ bands, thus providing the major new element of colour information for *Gaia* DR2 sources. The mean value of the G-band fluxes is reported for all sources while for about 80 per cent of the sources the mean values of the $G_{\rm BP}$ and $G_{\rm RP}$ fluxes are provided (for a small fraction of these sources only the $G_{\rm RP}$ value is reported). The photometric data processing considered three types of sources, "Gold", "Silver", and "Bronze", which represent decreasing quality levels of the photometric calibration achieved, where in the case of the Bronze sources no colour information is available. The photometric nature of each source is indicated in the released catalogue by a numeric field (`phot_proc_mode`) assuming values 0, 1 and 2 for gold, silver, and bronze sources respectively. At the bright end the photometric uncertainties are dominated by calibration effects which are estimated to contribute 2, 5, and 3 mmag RMS per CCD observation, respectively for G, $G_{\rm BP}$, and $G_{\rm RP}$ (Evans et al. 2018). For details on the photometric processing and the validation of the results we refer to Riello et al. (2018) and Evans et al. (2018).

The broad-band colour information suffers from strong systematic effects at the faint end of the survey ($G \gtrsim 19$), in crowded regions, and near bright stars. In these cases the photometric measurements from the blue and red photometers suffer from an insufficiently accurate background estimation and from the lack of specific treatment of the prism spectra in crowded regions, where the overlapping of images of nearby sources is not yet accounted for. This leads to measured fluxes that are inconsistent between the G and the $G_{\rm BP}$ and $G_{\rm RP}$ bands in the sense that the sum of the flux values in the latter two bands may be significantly larger than that in G (whereas it is expected that for normal spectral energy distributions the sum of fluxes in $G_{\rm BP}$ and $G_{\rm RP}$ should be comparable to that in G). A quantitative indication of this effect is included in *Gaia* DR2 in the form of the "flux excess factor" (the `phot_bp_rp_excess_factor` field in the data archive).

The distribution of the astrometric and photometric data sets in G is shown in purple in Fig. 1, where for comparison the distribution for *Gaia* DR1 is also shown in yellow. Note the improved completeness at the bright end of the survey and the improved photometry (less extremely faint sources) and completeness at the faint end. The distribution of the *Gaia*-CRF2 sources (pink-red line) shows a sharp drop at $G = 21$ which is because only QSOs at $G < 21$ were used for the construction of the reference frame.

3.3. Radial velocity data set

The radial velocity data set contains the median radial velocities, averaged over the 22 month time span of the observations, for 7 224 631 sources which are nominally brighter than 12th magnitude in the $G_{\rm RVS}$ photometric band. For the selection of sources to process, the provisional $G_{\rm RVS}$ magnitude as listed in the Initial Gaia Source List (Smart & Nicastro 2014) was used. The actual magnitudes in the $G_{\rm RVS}$ band differ from these provisional values, meaning that the magnitude limit in $G_{\rm RVS}$ is not sharply defined. In practice the sources for which a median radial velocity is listed mostly have magnitudes brighter than 13 in G (see light green line in Fig. 1). The signal to noise ratio of the RVS spectra depends primarily on $G_{\rm RVS}$, which is not listed in *Gaia* DR2. It was decided not to publish the $G_{\rm RVS}$ magnitude in *Gaia* DR2 because the processing of RVS data was focused on the production of the radial velocities, and the calibrations necessary for the estimation of the flux in the RVS passband (background light corrections and the knowledge of the PSF in the direction perpendicular to *Gaia*'s scanning direction) were only preliminary. As a result the $G_{\rm RVS}$ magnitudes were of insufficient quality for publication in *Gaia* DR2 (Sartoretti et al. 2018). The value of $G_{\rm RVS}$ as determined during the data processing was however used to filter out stars considered too faint ($G_{\rm RVS} > 14$) for inclusion in the radial velocity data set. For convenience we provide here a relation which allows to predict the value of $G_{\rm RVS}$ from the $(G - G_{\rm RP})$ colour.

$$G_{\rm RVS} - G_{\rm RP} = 0.042319 - 0.65124(G - G_{\rm RP}) + 1.0215(G - G_{\rm RP})^2 - 1.3947(G - G_{\rm RP})^3 + 0.53768(G - G_{\rm RP})^4$$
$$\text{to within 0.086 mag RMS for } 0.1 < (G - G_{\rm RP}) < 1.4\,, \quad (2)$$

and

$$G_{\rm RVS} - G_{\rm RP} = 132.32 - 377.28(G - G_{\rm RP}) + 402.32(G - G_{\rm RP})^2 - 190.97(G - G_{\rm RP})^3 + 34.026(G - G_{\rm RP})^4$$
$$\text{to within 0.088 mag RMS for } 1.4 \leq (G - G_{\rm RP}) < 1.7\,. \quad (3)$$

This relation was derived from a sample of stars for which the flux in the RVS band could be determined to a precision of 0.1 mag or better.

Radial velocities are only reported for stars with effective temperatures in the range 3550–6900 K (where these temperatures refer to the spectral template used in the processing, not to the $T_{\rm eff}$ values reported as part of the astrophysical parameter data set). The uncertainties of the radial velocities are summarised in Table 3. At the faint end the uncertainties show a dependency on stellar effective temperature, where the values are approximately 1.4 km s^{-1} and 3.6 km s^{-1} at $G_{\rm RVS} = 11.75$ for stars with $T_{\rm eff} \sim 5500$ K and $T_{\rm eff} \sim 6500$ K, respectively. The distribution over G of the sources with radial velocities shown in Fig. 1 in light green reflects the fact that over the range $4 < G < 12$ the completeness of the radial velocity data set with respect to the *Gaia* DR2 data set varies from 60 to 80 per cent (Katz et al. 2018). At the faint end ($G > 13$) the shape of the distribution is determined by the selection of stars for which radial velocities were derived (using the provisional value of $G_{\rm RVS}$) and the large differences between G and $G_{\rm RVS}$ that can occur depending on the effective temperature of the stars. For the details on the radial velocity data processing and the properties and validation of the resulting radial velocity catalogue we refer to Sartoretti et al. (2018) and Katz et al. (2018). The set of standard stars that was used to define the zeropoint of the RVS radial velocities is described in Soubiran et al. (2018).

3.4. Variability data set

The variability data set consists of 550 737 sources that are securely identified as variable (based on at least two transits of the sources across the fields of view of the two *Gaia* telescopes) and for which the photometric time series and corresponding statistics are provided. This number still represents

Table 3. Basic performance statistics for *Gaia* DR2.

Data product or source type	Typical uncertainty
Five-parameter astrometry (position & parallax)	0.02–0.04 mas at $G < 15$ 0.1 mas at $G = 17$ 0.7 mas at $G = 20$ 2 mas at $G = 21$
Five-parameter astrometry (proper motion)	0.07 mas yr^{-1} at $G < 15$ 0.2 mas yr^{-1} at $G = 17$ 1.2 mas yr^{-1} at $G = 20$ 3 mas yr^{-1} at $G = 21$
Two-parameter astrometry (position only)	1–4 mas
Systematic astrometric errors (averaged over the sky)	<0.1 mas
Gaia-CRF2 alignment with ICRF	0.02 mas at $G = 19$
Gaia-CRF2 rotation with respect to ICRF	<0.02 mas yr^{-1} at $G = 19$
Gaia-CRF2 alignment with ICRF	0.3 mas at $G < 12$
Gaia-CRF2 rotation with respect to ICRF	<0.15 mas yr^{-1} at $G < 12$
Mean G-band photometry	0.3 mmag at $G < 13$ 2 mmag at $G = 17$ 10 mmag at $G = 20$
Mean G_{BP}- and G_{RP}-band photometry	2 mmag at $G < 13$ 10 mmag at $G = 17$ 200 mmag at $G = 20$
Median radial velocity over 22 months	0.3 km s^{-1} at $G_{RVS} < 8$ 0.6 km s^{-1} at $G_{RVS} = 10$ 1.8 km s^{-1} at $G_{RVS} = 11.75$
Systematic radial velocity errors	<0.1 km s^{-1} at $G_{RVS} < 9$ 0.5 km s^{-1} at $G_{RVS} = 11.75$
Effective temperature T_{eff}	324 K
Extinction A_G	0.46 mag
Colour excess $E(G_{BP} - G_{RP})$	0.23 mag
Radius	10%
Luminosity	15%
Solar system object epoch astrometry	1 mas (in scan direction)

Notes. The astrometric uncertainties as well as the Gaia-CRF2 alignment and rotation limits refer to epoch J2015.5 TCB. The uncertainties on the photometry refer to the mean magnitudes listed in the main Gaia DR2 catalogue.

only a small subset of the total amount of variables expected in the *Gaia* survey and subsequent data releases will contain increasing numbers of variable sources. Of the sources identified as variable 363 969 were classified into one of nine variable types by a supervised light curve classifier. The types listed in the *Gaia* DR2 are: RR Lyrae (anomalous RRd, RRd, RRab, RRc); long period variables (Mira type and Semi-Regulars); Cepheids (anomalous Cepheids, classical Cepheids, type-II Cepheids); δ Scuti and SX Phoenicis stars. A second subset of 390 529 variable stars (largely overlapping with the variability type subset) was analysed in detail when at least 12 points were available for the light curve. These so-called "specific object studies" (SOS) were carried out for variables of the type Cepheid and RR Lyrae, long period variables, short time scale variables (with brightness variations on time scales of one day or less), and rotational modulation variables.

Figure 1 shows in dark blue the distribution over G of the sources identified as variable. The mean G value as determined in the photometric data processing (used in Fig. 1) may differ from the mean magnitude determined from the photometric time series where the variable nature of the source is properly accounted for. Hence the distribution in Fig. 1 should be taken as illustrative only. For full details on the variable star processing and results validation we refer to Holl et al. (2018) and references therein.

3.5. Astrophysical parameter data set

The astrophysical parameter data set consists of estimated values of T_{eff}, extinction A_G and reddening $E(G_{BP} - G_{RP})$ (both derived from the apparent dimming and reddening of a source), radius, and luminosity for stars brighter than $G = 17$. Table 1 contains the source counts for each of these astrophysical parameters. The magnitude distribution shown in Fig. 1 in cyan concerns all sources for which T_{eff} was estimated and indicates that this parameter is available for practically all sources at $G < 17$. Values of T_{eff} are only reported over the range 3000–10 000 K, which reflects the limits of the training data for the algorithm used to estimate T_{eff}. Estimates of the other astrophysical parameters are published for about 50% of the sources for which T_{eff} is published. This is caused by the filtering of the pipeline results to remove parameter estimates for which the input data are too

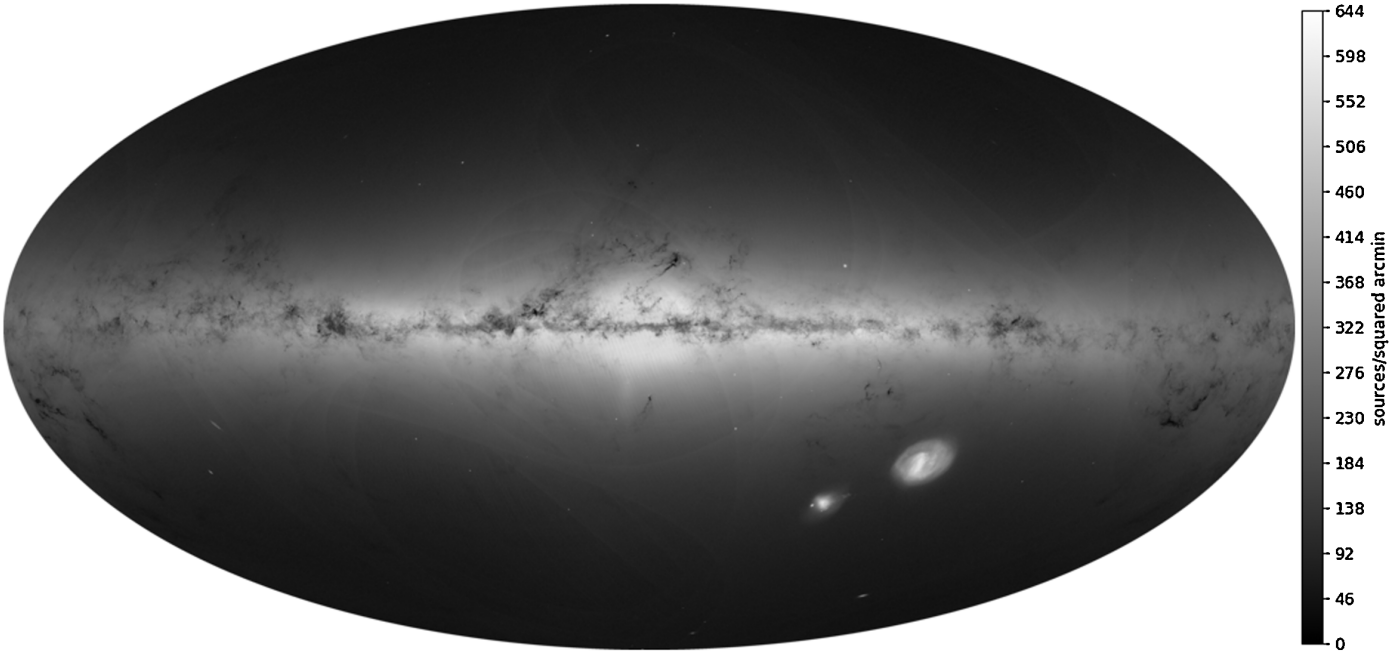

Fig. 3. Sky distribution of all *Gaia* DR2 sources in Galactic coordinates. This image and the one in Fig. 4 are Hammer projections of the full sky. This projection was chosen in order to have the same area per pixel (not strictly true because of pixel discretisation). Each pixel is ~5.9 square arcmin. The colour scale is logarithmic and represents the number of sources per square arcmin.

poor or for which the assumptions made lead to invalid results. The details of the astrophysical parameter processing and the validation of the results are described in Andrae et al. (2018).

3.6. Solar system objects data set

The solar system objects data set features epoch astrometry and photometry for a pre-selected list of 14 099 known minor bodies in the solar system, primarily main belt asteroids. Epoch astrometry refers to the fact that the measured celestial position for a given SSO is listed for each instance in time when it passed across the field of view of one of *Gaia*'s telescopes. The celestial positions at each epoch are given as seen from *Gaia*. These measurements can be used to determine orbits for the SSOs and the results thereof are described in Gaia Collaboration (2018f). For details on the processing of SSOs we refer to the same paper.Over the apparent magnitude range $G \sim 12$–17 the typical focal plane transit level of uncertainty achieved for the instantaneous SSO celestial positions is 1 mas in the *Gaia* scanning direction. Figure 1 shows in dark green the magnitude distribution for the SSOs, where it should be noted that the magnitudes as can be measured by *Gaia* represent instantaneous measurements taken far from opposition. Hence the magnitude histogram is to be taken as illustrative only.

4. Scientific performance and potential of *Gaia* DR2

Gaia DR2 is accompanied by six papers that provide basic demonstrations of the scientific quality of the results included in this release. The topics treated by the papers are:

- the reference frame *Gaia*-CRF2 (Gaia Collaboration 2018e);
- orbital fitting of the epoch astrometry for solar system objects (Gaia Collaboration 2018f);
- variable stars as seen in the *Gaia* DR2 colour-magnitude diagram (Gaia Collaboration 2018b), where the motion of variables in colour-magnitude space is explored;
- the kinematics of the Milky Way disk (Gaia Collaboration 2018d), illustrating in particular the power of having radial velocities available in *Gaia* DR2;
- the kinematics of globular clusters, the LMC and SMC, and other dwarf galaxies around the Milky Way (Gaia Collaboration 2018c), showcasing the power of *Gaia* DR2 to study distant samples of stars;
- the observational Hertzsprung-Russell diagram is explored in Gaia Collaboration (2018a).

We strongly encourage the reader to consult these papers for a full impression of the enormous scientific potential of the second *Gaia* data release.

Here we restrict ourselves to illustrating both the improvement in the data quality and the expanded set of data products through the updated map of the *Gaia* sky. Figure 3 shows the sky distribution of all the sources present in *Gaia* DR2 in the form of source densities on a logarithmic scale. When comparing to the map produced from *Gaia* DR1 data (Gaia Collaboration 2016a) it is immediately apparent that there is a strong reduction in the artefacts caused by the combination of source filtering and the *Gaia* scanning law (see Gaia Collaboration 2016a, for a more detailed explanation of these artefacts), which is another illustration of the increased survey completeness of *Gaia* DR2. Nonetheless there are still source count variations visible, which clearly are imprints from the scanning law (as executed over the first 22 months of the mission). For example there are two arcs above and below the ρ Oph clouds that can be traced all the way down to and below the Galactic plane (these can best be seen in the electronic version of the figure). Such arcs occur all along the ecliptic plane and are regions on the sky that were scanned more frequently by *Gaia* and therefore contain relatively more sources that were observed often enough for inclusion in the published catalogue.

One newly visible (and real) feature in this map is the Sagittarius dwarf which can be noted as an excess in star counts in a strip below the bulge region, stretching to the R Corona Australis region.

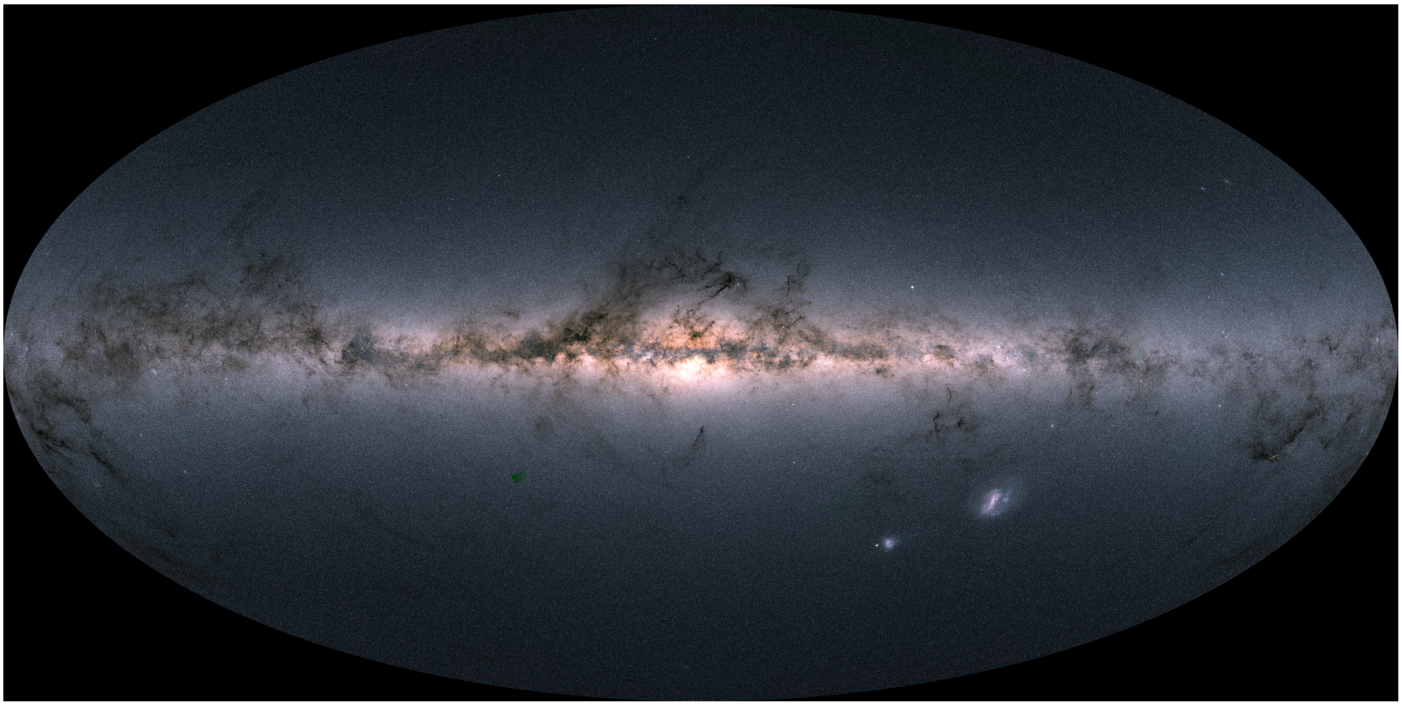

Fig. 4. Map of the total flux measured in the G_{RP}, G, and G_{BP} bands, where the flux in these bands is encoded in the red, green, and blue channel, respectively. There is one easily visible artefact in this map, a "green" patch to the lower left of the bulge which is a region where G_{BP} and G_{RP} data are not available for a large number of sources, leading to the greenish colour which was used to encode the G-band fluxes (which are available for all sources). Such artefacts also occur (although not as visible) in the region to the upper left of the Small Magellanic Cloud and at high Galactic latitude to the right of the north Galactic pole region. The areas where green patches are likely to occur can be identified in Fig. 27 in Evans et al. (2018) which shows the celestial distribution of *Gaia* DR2 sources for which no BP/RP photometry is available.

Figure 4 shows a map that combines the integrated fluxes as observed in the G_{RP}, G, and G_{BP} bands, where the integrated flux map for each of the bands was used to colour code the image according to a red, green, and blue channel. The map illustrates the availability of homogeneous all-sky multi-band photometry in *Gaia* DR2 and offers a magnificent view of the Milky Way in colour. This flux map also reveals numerous open clusters which are not readily visible in the source count map (while on the other hand many faint source concentrations, such as distant dwarf galaxies are no longer visible). Complete details on the construction of the images in Figs. 3 and 4 are provided in Moitinho et al. (2018).

One aspect of the sky maps shown in Figs. 3 and 4 that is perhaps not as well appreciated is their effective angular resolution, which given the size of *Gaia*'s main telescope mirrors (1.45 m along the scanning direction, Gaia Collaboration 2016b) should be comparable to that of the *Hubble* Space Telescope. Gaia Collaboration (2016a) and Arenou et al. (2017) discuss how the effective angular resolution of *Gaia* DR1 is limited to about 2–4 arcsec owing to limitations in the data processing. This has much improved for *Gaia* DR2. The gain in angular resolution is illustrated in Fig. 5. The top panel shows the distribution of source pair distances in a small, dense field. For *Gaia* DR2 (upper, red curve) source pairs below 0.4–0.5 arcsec are rarely resolved, but the resolution improves rapidly and above 2.2 arcsec practically all pairs are resolved. For *Gaia* DR1 the fraction of resolved source pairs started to fail at separations of 3.5 arcsec, reaching very low values below 2.0 arcsec. The same, modest resolution is seen for *Gaia* DR2 if we only consider sources with G_{BP} and G_{RP} photometry. The reason is the angular extent of the prism spectra and the fact that *Gaia* DR1 only includes sources for which the integrated flux from the BP/RP spectra could be reliably determined. The lower panel shows in the same way the source pairs in the one hundred times larger, sparse field. The more remarkable feature here is the peak of resolved binaries at small separations, which was missed in *Gaia* DR1. A similar population must be present in the dense field, where it cannot be discerned because the field is dominated by distant sources. The figure also demonstrates that the gain in number of sources from *Gaia* DR1 to *Gaia* DR2 is mainly due to the close source pairs. Finally, Fig. 5 clearly demonstrates that the effective angular resolution of *Gaia* DR2 quite significantly exceeds that of all ground-based large-area optical sky surveys.

5. Treat *Gaia* DR2 as independent from *Gaia* DR1

Although *Gaia* DR1 and *Gaia* DR2 are based on observations from the same instruments, the discussion in the following subsections shows that the two releases should be treated as independent. In particular the tracing of sources from *Gaia* DR1 to *Gaia* DR2 (should this be needed for a particular application) must be done with care.

5.1. Gaia DR2 represents a stand-alone astrometric catalogue

Because the observational time baseline for *Gaia* DR2 is sufficiently long, parallax and proper motion can be derived from the *Gaia* observations alone. That is, the *Tycho-Gaia* Astrometric Solution (TGAS, Michalik et al. 2015a) as employed for the 2 million brightest stars in *Gaia* DR1 is no longer needed, and the astrometric results reported in *Gaia* DR2 are based solely on *Gaia* observations. For the TGAS subset from *Gaia* DR1 there is thus a large difference in the time baseline for the proper motions